Accuracy of Geophysical Parameters Derived from AIRS/AMSU as a Function of Fractional Cloud Cover

Joel Susskind[a], Chris Barnet[b], John Blaisdell[c], Lena Iredell[c], Fricky Keita[c], Lou Kouvaris[c], Gyula Molnar[d], and Moustafa Chahine[e]
[a]NASA Goddard Space Flight Center, Greenbelt, MD, USA 20771
[b]NOAA/NESDIS, Camp Springs, MD, USA 20746
[c]Science Applications International Corporation, NASA GSFC, Greenbelt, MD, USA 20771
[d]Joint Center for Earth Systems Technology, NASA GSFC, Greenbelt, MD, USA 20771
[e]Jet Propulsion Laboratory, Pasadena, CA, USA 94709

Abstract. AIRS was launched on EOS Aqua on May 4, 2002, together with AMSU A and HSB, to form a next generation polar orbiting infrared and microwave atmospheric sounding system. The primary products of AIRS/AMSU are twice daily global fields of atmospheric temperature-humidity profiles, ozone profiles, sea/land surface skin temperature, and cloud related parameters including OLR. The sounding goals of AIRS are to produce 1 km tropospheric layer mean temperatures with an rms error of 1K, and layer precipitable water with an rms error of 20%, in cases with up to 80% effective cloud cover. The basic theory used to analyze AIRS/AMSU/HSB data in the presence of clouds, called the at-launch algorithm, was described previously. Pre-launch simulation studies using this algorithm indicated that these results should be achievable. Some modifications have been made to the at-launch retrieval algorithm as described in this paper. Sample fields of parameters retrieved from AIRS/AMSU/HSB data are presented and validated as a function of retrieved fractional cloud cover. As in simulation, the degradation of retrieval accuracy with increasing cloud cover is small. HSB failed in February 2005, and consequently HSB channel radiances are not used in the results shown in this paper. The AIRS/AMSU retrieval algorithm described in this paper, called Version 4, become operational at the Goddard DAAC in April 2005 and is being used to analyze near-real time AIRS/AMSU data. Historical AIRS/AMSU data, going backwards from March 2005 through September 2002, is also being analyzed by the DAAC using the Version 4 algorithm.

1. Introduction

AIRS/AMSU/HSB is a state of the art advanced infra-red microwave sounding system that was launched on the EOS Aqua platform in a 1:30 AM/PM sun synchronous orbit on May 4, 2002. An overview of the AIRS instrument and the objectives for AIRS/AMSU/HSB is a given in Aumann et al. (2003). The sounding goals of AIRS are to produce 1 km tropospheric layer mean temperatures with an rms error of 1K, and layer precipitable water with an rms error of 20%, in cases with up to 80% effective cloud cover. Aside from being part of a climate mission, one of the objectives of AIRS is to provide sounding information of sufficient accuracy such that when assimilated into a general circulation model, significant improvement in forecast skill would arise. The at-launch algorithm to produce level 2 products (geophysical parameters) using AIRS/AMSU/HSB data, and expected results based on simulation studies, are given in Susskind et al. (2003). The results of that simulation indicate that the sounding goals of AIRS/AMSU/HSB should be achievable. In that simulation, perfect knowledge of the instrumental spectral response functions and the inherent physics of the radiative transfer equations was assumed. Therefore, if the true state of the atmosphere and underlying surface were known perfectly, one could compute the radiances AIRS, AMSU, and HSB would see exactly up to instrumental noise. Susskind et al. (2003) alluded to the fact that this is not the case in reality, and additional terms would have to be included in the retrieval algorithm to account for systematic differences (biases) between observed brightness temperatures and those computed knowing the "true" surface and atmospheric state, and also to account for residual computational errors after that systematic bias is accounted for (computational noise).

In this paper, we show results based on the algorithm we were using to analyze AIRS/AMSU data on November 30, 2004, which we will refer to as Version 4. This algorithm is very similar to the at-launch version, with the major differences attributable to the factors described above, as well as changes in the internally generated quality flags. As in Susskind et al. (2003), the retrieval methodology, including quality control, is based only on AIRS and AMSU observations, and does not involve use of a GCM forecast model except for its use to provide the surface pressure. HSB failed in February 2004, and HSB radiances are not used in the current retrieval algorithm so as to allow for a continuous climate data record before and after the loss of HSB. Loss of HSB did not appreciably affect the quality of the retrieved data. AMSU A channel 7 was also found to be very noisy and is not included in any of the calculations.

The Goddard DAAC began processing AIRS/AMSU data using Version 4 in April 2005. JPL delivered an earlier version of the algorithm to the Goddard DAAC, Version 3, for the earliest near real time processing of AIRS level 2 products starting in July 2003. In this paper, we outline the differences between the at-launch version and Version 4, and show sample results using Version 4 on data for September 6, 2002, with a particular emphasis on the quality of the retrieved parameters as a function of fractional cloud cover. Research to further improve the results of analysis of AIRS/AMSU data, leading to a Version 5 algorithm, is continuing.

2. Differences Between the At-Launch Algorithm and Version 4

The differences between Version 4 and the at-launch version of the retrieval algorithm are relatively small. The post-launch channel frequencies were somewhat different from those pre-launch, as expected, as were the channel spectral response functions. Consequently, new Radiative Transfer Algorithm (RTA) coefficients were generated (Strow et al., 2005) to be consistent with the post-launch instrumental conditions. Minor modifications were therefore made to the set of channels used in the retrieval algorithm. The most significant of these resulted from the finding that more channels in the 4.3 μm region were affected by non-local thermodynamic equilibrium (non-LTE) than previously thought. Radiances in these channels are perturbed during the day, and these channels are currently not used in the retrieval algorithm day or night. It was also found that observed channel brightness temperatures for AIRS, as well as AMSU, were biased from those computed using the RTA with the best estimate of the truth. These biases, referred to as "tuning coefficients", are subtracted from all terms in the retrieval algorithm involving observed minus computed brightness temperatures. New regression coefficients were also generated (Goldberg et al., 2005) based on clear column radiances for an ensemble of accepted retrievals, using the 3-hour ECMWF forecast at "truth". A few AIRS channels exhibit a radiometric instability characteristic, known as "popping", and these channels are excluded from the list of channels used either in the regression or the physical retrieval steps. It was also found that many of the channels used in the at-launch physical retrieval algorithm were not needed in practice, and are no longer used in the physical retrieval steps so as to make the physical retrieval computationally more efficient with no loss of accuracy. A new concept has also been introduced in terms of quality control, in which

different geophysical parameters retrieved from AIRS/AMSU data have different criteria for acceptance.

Details of the at-launch AIRS/AMSU retrieval algorithm are given in Susskind et al. (2003). The basic steps in the retrieval algorithm are identical to those shown in Susskind et al. (2003). The only change is a new step to determine the CO profile done after the retrieval of the final surface skin temperature and temperature profile. This step is analogous to what is done in the H_2O and O_3 profile retrieval steps. In this paper areas in which additional changes to the methodology of Susskind et al. (2003) were made will be briefly summarized. More details will be published separately in a revised AIRS Algorithm Theoretical Basis Document (ATBD). The major change from the at-launch algorithm is with regard to the new quality flag concept. This will be described in detail.

2.1 Generation of Tuning Coefficients

Steps in the physical retrieval and cloud clearing algorithms involve the difference between observed (cloud cleared) radiances $\hat{R}_i$ and those computed from some geophysical state, R_i^{comp} using the radiative transfer algorithm (RTA) described in Strow et al. (2005). If one had a perfectly calibrated instrument and perfect parameterization of the radiative transfer physics, then, given the true surface and atmospheric state, the observed radiances $\hat{R}_i$ could be calculated up to instrumental noise. Systematic errors in either the calibration of the observed radiances $\hat{R}_{i,\ell}$ (channel i, zenith angle ℓ), or in the computation of radiances $R_{i,\ell}^{comp}$, would introduce biases in $\left(\hat{R}_{i,\ell} - R_{i,\ell}^{comp}\right)$ and propagate errors into the solution. We attempt to identify these biases and remove their effect by subtracting them from all terms $\left(\hat{R}_i - R_i^{comp}\right)$ whenever they occur in the retrieval and cloud

clearing processes. This subtraction is done in the brightness temperature domain for both AIRS and AMSU radiances, in a manner analogous to that described in Susskind and Pfaendtner (1989) and used by Susskind et al. (1997) in the analysis of HIRS2 and MSU radiances:

$$\left(\hat{\Theta}_{i,\ell} - \Theta_{i,\ell}^{comp}\right)' = \left(\hat{\Theta}_{i,\ell} - \Theta_{i,\ell}^{comp}\right) - \delta\Theta_{i,\ell} \qquad 1)$$

where Θ_i is the brightness temperature corresponding to R_i and $\delta\Theta_{i,l}$ is the tuning correction. All retrieval steps involve $\hat{R}_i - R_i^{comp}$. For AIRS channels, the tuned value of $\left(\hat{R}_{i,\ell} - R_{i,\ell}^{comp}\right)$ is given by

$$\left(\hat{R}_{i,\ell} - R_{i,\ell}^{comp}\right)' = \left(\hat{\Theta}_{i,\ell} - \Theta_{i,\ell}^{comp}\right)' \left(\frac{dB}{dT}\right)_{\hat{\Theta}_{i,\ell}}, \qquad 2)$$

and is used in place of $\left(\hat{R}_{i,\ell} - R_{i,\ell}^{comp}\right)$ in all retrieval steps. An analogous procedure is used to adjust observed minus computed radiances in the cloud clearing step.

The methodology used to generate the tuning coefficients is to identify systematic differences between $\hat{R}_{i,l}$ and $R_{i,l}^{true}$, where $R_{i,\ell}^{true}$ is the radiance computed using the "true" geophysical conditions for channel i and earth location corresponding to zenith angle ℓ. In order to generate AIRS channel tuning coefficients, cases were selected thought to be unaffected by clouds so as not to have to account for cloud effects on the observed radiances. The 3-hour ECMWF forecast, collocated to the satellite observations, is used as truth, and observations were limited to nighttime non-frozen ocean (henceforth referred to as "ocean") so as to avoid effects of solar radiation reflected by the surface as well as effects of non-LTE. Ocean night cases on September 6, 2002 were used to determine the biases. These biases had very little scene or zenith angle

dependence. Therefore, for AIRS channels, the tuning coefficient for channel i is taken as a constant

$$\delta\Theta_{i,\ell} = A_i. \qquad 3)$$

where A_i is the average bias over all scenes and zenith angles. AIRS tuning coefficients, determined in this manner, are applied only to channels in the CO_2 and N_2O absorption regions.

The procedure used to generate AMSU tuning coefficients is analogous to that used to generate AIRS tuning coefficients. The coefficients used were generated for ocean cases on September 6, 2002. These cases were screened to eliminate contamination from precipitating clouds. Unlike the biases found for AIRS channels, AMSU channels had a pronounced, and systematic, zenith angle (beam position) dependence. This arises from effects of antenna side-lobes, which were not adequately accounted for in the calibration of the AMSU observations. Based on this finding, AMSU channels are tuned according to

$$\delta\Theta_{i,\ell} = A_{i,\ell} \qquad 4)$$

with $A_{i,\ell}$ determined using the September 6, 2002 data. The AIRS and AMSU tuning coefficients, determined using ocean cases on September 6, 2002, are used globally for all time periods.

An empirical estimate of the uncertainty of the tuning coefficients was also derived in a manner described in the AIRS ATBD. This uncertainty, generally on the order of 0.3K, is added to the diagonal term of the channel noise covariance matrix.

2.2 Other Minor Differences from the At-Launch Version

The basic algorithms for cloud clearing and retrieval of geophysical parameters are identical to those described in Susskind et al. (2003). A few minor details have been modified, primarily with

regard to the number of channels used in each retrieval step. In general, less channels are used in the physical retrieval steps so as to decrease the computational time required to perform the physical retrievals, with little or no change in accuracy. Some modifications have also been made to the damping parameters ΔB_{max} used in each retrieval step. In addition, a CO profile retrieval step is now included in the retrieval process which is totally analogous to the H_2O and O_3 profile retrieval steps. Details of these modifications are given in the AIRS ATBD.

3 New Quality Flags

The major change to the at-launch algorithm is a new concept with regard to quality flags. Susskind et al. (2003) discussed a number of threshold tests used to determine whether the combined IR/MW retrieval is of good quality. These tests utilize only the AIRS/AMSU radiance data. No external data, such as GCM forecast fields or MODIS observations are used. If the tests were all passed, the combined IR/MW retrieval state, and associated clear column radiances, were reported, as well as cloud and OLR values consistent with the AIRS radiance observations and the IR/MW retrieval state. If any of the tests were not passed, IR/MW retrieval state was "rejected" and the MW/strat IR retrieval state was reported, as well as associated values of cloud parameters and OLR constant with that state. Rejection usually implied problems with regard to treating effects of clouds in the field of view, and rejected cases produced generally poorer results in the mid-lower troposphere and at the surface.

The basic approach used now with regard to quality flags is identical with one major exception: different quality flags are used for different geophysical parameters. Problems dealing with clouds in the field of regard (3x3 array of AIRS fields of view) may produce a poor temperature profile in the lower troposphere, but should not degrade accuracy of the stratospheric temperature or upper

tropospheric water vapor. For this reason, a less strict threshold test is applied to accept stratospheric temperatures than lower tropospheric temperatures. Cases are classified 0-6 according to their ability to pass 6 increasingly more stringent threshold tests. The higher the number, the tighter the test which is passed. Class 6 passes the tight sea surface temperature test, Class 5 passes the standard sea surface temperature test, Class 4 passes the lower tropospheric temperature test, Class 3 passes the mid-tropospheric temperature test, Class 2 passes the constituent profile test, Class 1 passes the stratospheric temperature test, and Class 0 fails the stratospheric temperature test. The final IR/MW retrieval state and associated clear column radiances and cloud and OLR fields are provided for all Classes except for 0, in which case the MW/strat IR state and associated cloud and OLR parameters are reported.

The threshold tests used to assign quality flags are for the most part identical to those in Susskind et al. (2003), with the addition of 4 new tests. As before, all tests involve only AIRS and AMSU radiances. Susskind et al. (2003) tested for the following quantities: retrieved effective cloud fraction $\alpha\varepsilon$; retrieved total liquid water content W_{liq}; the RMS difference between the temperature profile in the lowest 3 km of the atmosphere retrieved in the final state and the test microwave only retrieval, $\Delta T(P)$; the final cloud noise amplification factor, $A^{(4)}$; the final effective cloud noise amplification factor, $A_{eff}^{(4)}$; the RMS value of the weighted residuals of the clear column brightness temperatures of the channels used in the cloud clearing process, ΔF; the ratio of the weighted residuals of the channels used to determine the final temperature profile to its theoretical value, R_{temp}; and the analogous ratio for the channels used to determine the final surface parameters, R_{surf}. Definitions of all quantities referred to above are given in Susskind et al. (2003) and will not be repeated here. Susskind et al. (2003) threshold values for all of these tests are shown

in the first column of Table 1. A test is passed if the value of the parameter used in the test is less than or equal to the threshold value. All tests must be passed for the final IR/MW retrieval state to be accepted.

Version 4 threshold values for each of the 6 Classes described above for all of these tests are shown in Table 1. Tests for some classes use separate threshold values for ocean cases and land cases. When the thresholds are different, the land threshold is shown in parenthesis, and is always larger or the test is non-applicable. Non-applicable tests are indicated by X. Four new tests have also been added: $A_{eff}^{(1)}$, which is analogous to $A_{eff}^{(4)}$ but is applied after the initial cloud clearing; $\Delta\Theta_5$, which is the absolute value of the (tuned) difference between the observed brightness temperature of AMSU channel 5 and that computed from the final retrieval state; Δ_{tskin}, which is the absolute value of the difference between the regression surface skin temperature and the final surface skin temperature; and RS (Goldberg et al., 2003), which represents how well the observed AIRS radiances can be represented by use of 200 principle components. Threshold values for these tests for all Classes are included in Table 1. Bold values in a class indicate the introduction of a new test or tightening of a previous threshold.

3.1 The stratospheric temperature test

This is the most fundamental test and is used to indicate, first and foremost, whether the final combined IR/MW retrieval, including associated clear column radiances and cloud and OLR parameters should be used, or whether the combined IR/MW retrieval should be "rejected" in all its aspects. The IR/MW retrieval is "rejected" if it is thought to be poorer than the MW/strat IR retrieval, which uses no AIRS channels affected by clouds. The combined IR/MW retrieval cannot always be used because cloud clearing cannot be done under overcast conditions. If the final

retrieval were used under such conditions, not only would very poor (too cold) tropospheric and surface skin conditions be derived, but using those conditions to determine cloud fields would result in little or no fractional cloud cover being derived, because AIRS channel radiances computed using the retrieved state would match observed radiances, without the need to add clouds to the scene (see Susskind et al., 2003). Products derived from the combined final IR/MW retrieval are rejected if the retrieved effective cloud fraction is 90% or more. Two tests are added to make sure the clear column radiances are acceptable: $A_{eff}^{(1)}$ must be less than 200 and RS must be less than 10. Failure of the first test indicates that the initial cloud clearing step had significant problems (note the $A_{eff}^{(4)}$ threshold was set equal to 8 in Susskind et al. (2003)) and of the second test indicates a significant problem with the observed AIRS radiances (RS equal to 1 is the expected value for nominal radiance performance). Retrieved temperatures 200 mb and above (lower pressures) are flagged as good if this test is passed.

3.2 The constituent profile test

This test is designed to insure that constituent profiles (O_3, CO, H_2O) are of sufficient accuracy for research use. Constituent profiles are considerably more variable, and less well predicted by models, than are temperature profiles. In general, the more spatial coverage one has, the better, provided the accuracy is adequate. This is especially true with regard to studying interannual variability of monthly mean differences. This applies particularly to water vapor, for which it is desirable to avoid a clear (dry) bias in the selection of the cases to be included in generation of the monthly mean fields. Most CO and H_2O exists in the troposphere, however, and ability to treat cloud effects on the radiances accurately is more important than with regard to

stratospheric temperatures. Therefore three tests used in Susskind et al. (2003), designed to indicate potential cloud clearing problems, are included in the constituent profile test as shown in Table 1. The liquid water test threshold is the same as is Susskind et al. (2003), and the thresholds for $A^{(4)}$ and ΔF are considerably less stringent.

3.3 The mid-tropospheric temperature test

Retrieved mid-tropospheric temperatures are affected more by errors in the treatment of clouds in the field of view than are stratospheric temperatures. Therefore, tighter quality control is employed in the mid-tropospheric temperature test. Susskind and Atlas (2004) showed that assimilation of AIRS temperature profiles retrieved from AIRS data, using an earlier version of the AIRS retrieval system (which employed a single rejection threshold for all geophysical parameters), significantly improved forecast skill. Moreover, the improvement was much larger if all accepted cases were used as opposed to use of the slightly more accurate, but much less frequent, temperature soundings in cases found to be clear. Therefore, from the data assimilation perspective, there is a trade-off between accuracy and spatial coverage, as is also true with regard to the study of interannual variability. The thresholds shown in Table 1 are designed to maximize spatial coverage, while minimizing loss in accuracy. Four tests used in Susskind et al. (2003) are are now included in the mid-tropospheric temperature test. The first test, $\Delta T(P)$, which contains the difference in the retrieved temperature in the lowest 3 km between the combined IR/MW retrieval and the test MW retrieval, is looser than that in Susskind et al. (2003). In addition, the threshold for ΔF has been tightened from that of the constituent profile test, but is still less stringent than in Susskind et al. (2003). Thresholds in the three additional new tests, $A^{(4)}$, R_{temp}, and R_{surf} are all somewhat

tighter than in Susskind et al. (2003). Thresholds for $A_{eff}^{(1)}$ and RS have also been tightened from their values in the constituent test, but are still at moderate values. A new test, $\Delta\Theta_5$ has also been added, requiring that the observed brightness temperature for AMSU channel 5, sensitive to lower tropospheric temperatures, should agree with that computed from the combined IR/MW retrieval to within 2K after tuning is applied. The threshold for ΔF over land is less restrictive than over ocean because ΔF is affected by uncertainty in surface emissivity, which is greater over land than over ocean. R_{surf} and $A_{eff}^{(1)}$ are also affected significantly by uncertainty in surface emissivity and for this reason, these tests are not utilized over land, so as to maximize spatial coverage. Errors in surface emissivity do not degrade retrieved mid-tropospheric temperatures appreciably. If the mid-tropospheric temperature test is passed, the temperature profile is flagged as good above 3 km of the surface.

3.4 The lower-tropospheric temperature test

Retrieved temperatures in the lowest 3 km of the atmosphere are most sensitive to cloud clearing errors, as well as errors in surface emissivity. A_{eff} and ΔF are both measures of how well cloud clearing is being done and potential problems with surface emissivity. The threshold for ΔF is now tightened considerably and is tighter than in Susskind et al. (2003), in which it had to be relaxed as a compromise so as not to reject the entire profile too often. $A_{eff}^{(1)}$ is also now used over land, and together with ΔF, flags many cases over arid land (in which retrieved surface emissivity can have large errors) as bad. The test Δ_{tskin} is also introduced which indicates a potential problem with the retrieved surface skin temperature.

3.5 Standard and tight sea surface temperature tests

Sea surface temperature is determined quite well by other instruments such as MODIS. Therefore, for AIRS to produce a useful sea surface temperature product for climate research, it must have very tight quality control. Surface skin temperature is also the product most affected by errors in the cloud clearing process, especially with regard to very low clouds. In the standard SST Test, thresholds for four tests have been tightened as shown in Table 1. This test is applied only over ocean, as land temperatures are less well measured by other instruments. The test most correlated with sea-surface temperature accuracy was $A_{eff}^{(1)}$, with lower values indicating more accurate sea-surface temperatures. The percent of accepted sea surface temperatures drops rapidly with decreasing acceptance thresholds however. If $A_{eff}^{(1)}$ is less than a second threshold, shown for Class 6, then the Tight SST Test is passed.

4. Results Using Version 4

One of the objectives of this paper is to show the extent that high quality soundings and clear column radiances are derived from AIRS/AMSU observations in the presence of clouds. The cloud clearing process does introduce noise in the derived cloud cleared radiances (Susskind et al., 2003). Therefore, one would expect a degradation in retrieval accuracy with increasing cloud cover. It is critical that this degradation should not be appreciable if the retrieved parameters are to be useful for weather and climate research purposes. In this paper, the accuracy of global geophysical parameters derived from AIRS/AMSU observations on September 6, 2002 was evaluated by comparison with a co-located ECMWF 3 hour forecast. The ECMWF forecast has errors of its own, and this should be

borne in mind when interpreting the results of the comparisons. Instead of an assessment of the absolute accuracy of the retrieved quantities, we concentrate on the degree of degradation in "accuracy", as defined by agreement with ECMWF, occurring with increasing cloud cover. Errors in the ECMWF "truth" may decrease the apparent differences in accuracy between clear and cloudy cases, but only by making the clear cases appear less accurate than they actually are, and not by making the cloudy cases appear more accurate than they are. In all cases, the quality control methodology described in Section 2.6 is used to include or exclude individual retrieved geophysical parameters from the figures shown.

Figure 1 shows the number of cases for each retrieved effective fractional cloud cover, in 0.5% bins, for the whole day September 6, 2002. The effective fractional cloud cover is given by the product of the fraction of the field of view covered by clouds and the cloud emissivity at 11 μm. The average global effective cloudiness was determined to be 42.61%. There are peaks at 0% and 100% effective cloud cover, with a very smooth distribution at intermediate effective cloud fraction. The discontinuity at 90% cloud cover is an artifact arising from the switch from clouds retrieved primarily using the IR/MW retrieved state to clouds retrieved using the MW/strat IR state. Also shown is the percent of accepted retrievals as a function of retrieved effective cloud cover for all cases passing the Stratospheric Temperature Test, the Constituent Test, the Mid-Tropospheric Temperature Test, and the Lower Tropospheric Temperature Test, as well as non-frozen ocean cases passing the standard SST Test and the Tight SST Test. Almost all cases with retrieved effective cloud fraction less than 90% pass the Stratospheric Temperature Test, with the percent yield falling slowly with increasing cloud cover, from close to 100% at low cloud fractions to about 65% at close to 90% cloud cover. 82% of the global cases pass the Stratospheric Temperature Test, with an average effective cloud fraction of 32.95%. 81% of the cases pass the slightly more restrictive

Constituent Test, with an average effective cloud fraction of 32.60%. 52% of the global cases pass the Mid-Tropospheric Temperature Test, with an acceptance rate of about 82% for low effective cloud fraction, falling to about 20% at 80% effective cloud fraction, and 10% at 90% effective cloud fraction. The previous acceptance methodology (Susskind et al., 2003) rejected all cases with effective cloud fraction greater than 80%. The mean effective cloud fraction for all cases passing the Mid-Tropospheric Temperature Test is 25.25%. Only 29% of the cases pass the Lower Tropospheric Temperature Test, primarily over ocean, with an acceptance rate near 55% for low cloud fractions falling to 7% at 80% effective cloud fraction and 3% at 90% effective cloud fraction, and an average effective cloud fraction of 20.40%. The SST acceptance tests are applied only over non-frozen ocean. The standard SST Test accepts about 21% of the ocean cases, with an acceptance rate of 50% under nearly clear conditions, and an average cloud fraction of 10.14%, while the Tight SST Test accepts only 9% of the cases, with an average effective cloud fraction of 7.05%. The Tight SST Test allows for many more cases than does the clear test (Susskind et al., 2003) which accepts only 3.69% of the cases.

Figure 2a shows the retrieved effective cloud top pressure and effective cloud fraction for ascending orbits on September 6, 2002. The global mean effective cloud fraction and its spatial standard deviation are indicated in the figure. The results are presented in terms of cloud fraction in 5 groups, 0-20%, 20-40%, etc. with darker colors indicating greater cloud cover. These groups are shown in each of 7 colors, indicative of cloud top pressure. The reds and purples indicate the highest clouds, and the yellows and oranges the lowest clouds. Cloud fields are retrieved for all cases in which valid AIRS/AMSU observations exist. Gray means no data was observed. Figure 2b shows the retrieved 200 mb temperature field. This demonstrates the coverage of cases where stratospheric temperatures are accepted. Gray indicates regions where either no valid observations

existed or the stratospheric temperature retrieval was rejected, generally in regions of cloud cover 90-100%. Figure 2c shows retrieved values of total precipitable water vapor. This demonstrates the coverage of constituent profiles. Figure 2d shows retrieval values of 500 mb temperature, demonstrating coverage of mid-tropospheric temperatures. Gaps in the data coverage of mid-tropospheric temperature due to extensive cloud cover are larger than for stratospheric temperatures. Retrieved fields are quite coherent, and show no apparent artifacts due to clouds in the field of view. Water vapor has considerably more fine scale structure than temperature and contains some very large spatial gradients. The extent of gaps in water vapor coverage due to areas of rejected retrievals (retrievals which fail the Constituent Test) are considerably less than with regard to the Mid-Tropospheric Temperature Test, but somewhat larger than with regard to the very loose Stratospheric Temperature Test. As shown in Figure 1, the percent of cases accepted as a function of increasing cloud cover for these two classes of retrievals is almost identical.

Figure 3a shows the difference between the retrieved 700 mb temperature and the ECMWF 6 hour forecast field for ascending orbits on September 6, 2002, for those cases passing the Lower Tropospheric Temperature Test, while Figure 3b shows the same field for all cases passing the looser Mid-Tropospheric Temperature Test. The difference in spatial coverage is significant, particularly over land where 700 mb temperature retrievals appear to be biased warm compared to the ECMWF forecast. Statistics showing the global mean difference from ECMWF, the spatial standard deviation of the difference, and the spatial correlation of the retrieved and ECMWF fields are included in the figures. The overall accuracy is somewhat better with the tighter acceptance criteria, and this difference is significant for data assimilation purposes. When statistics are shown depicting the accuracy of lower tropospheric temperatures (Figures 5-7), only cases passing the Lower Tropospheric Temperature Test are included. All data shown in Figure 3b is included in the

generation of monthly mean fields however, so as to allow for global coverage, especially over arid land regions.

Figures 3c and 3d shows the differences of retrieved ocean surface skin temperature (SST) from the ECMWF SST analysis for the ascending orbits of September 6, 2002. Figure 3c includes only those cases passing the Tight SST Test and Figure 3d also includes those cases passing the standard SST Test. A considerable increase in yield is obtained using the standard SST Test, with some degradation in accuracy of Sea Surface Temperatures. The biases compared to ECMWF are negative in both cases, with a larger negative bias found in cases passing the standard SST Test. Errors due to cloud clearing are typically negative, resulting from under-correcting for effects of clouds in the field of view. This would imply that the Tight SST Test is eliminating more cases where cloud clearing errors are resulting in poorer sea surface temperatures. Caution must be taken however because the ECMWF "truth" may have its own biases.

Figure 4 shows the number of daytime and nighttime ocean cases between 50°N and 50°S, on September 6, 2002, as a function of the difference of the retrieved SST from the ECMWF analysis. Results are shown for cases which passed the Tight SST Test, the standard SST Test, and the Lower Tropospheric Temperature Test. Figure 3c and 3d showed the spatial distribution differences for the daytime orbits applying each of the SST Tests. Figure 4 includes statistics compared to ECMWF, showing the percent of all oceanic cases 50°N-50°S included in the statistics, the mean difference, the standard deviation of the difference, and the percentage of outliers, defined as cases passing the test that differ from ECMWF by more than 3K. There is a small negative bias of retrieved Sea Surface Temperatures compared to ECMWF, that increases with increasing acceptance rate, from –0.21K for cases within the Tight SST Test, to – 0.76K for cases passing the Lower Tropospheric Temperature Test. The standard deviation of the cases from ECMWF also

increases slightly. On the other hand, the number of primarily cold outliers increases significantly, from 0.68% to 7.28%. Therefore, the Lower Tropospheric Temperature Test by itself is not adequate for the purpose of producing accurate monthly mean sea surface temperatures. As with all the test thresholds, experiments are being conducted to optimize the trade-off between spatial coverage and accuracy for best use in studying interannual monthly mean sea-surface temperature differences. Susskind et al. (2005) generated monthly mean surface skin temperature fields over ocean using the standard SST Test. Over land, retrievals passing the Mid-Tropospheric Temperature Test were used, so as to get good spatial coverage.

Figure 5a shows RMS differences from the ECMWF forecast of retrieved 1 km layer mean tropospheric temperatures, and 3 km layer mean stratospheric temperatures, for ocean cases on September 6, 2002. Results shown are for all cases passing the Stratospheric Temperature Test, the Mid-Tropospheric Temperature Test, the Lower Tropospheric Temperature Test, the standard SST Test, the Tight SST Test. Results for those cases passing an additional clear test, as defined by Susskind et al. (2003), are also included in the figure. The number of cases and percentage of all cases included in the statistics are indicated for each test.

Differences of retrieved stratospheric temperature from ECMWF are not appreciably different from one another for cases passing any of the quality flags. The large differences from ECMWF above 15 mb are primarily a result of the lower accuracy of the ECMWF "truth" in the upper stratosphere. Tropospheric soundings passing either of the tropospheric quality control tests agree with the ECMWF forecast on the order of 1K. Part of this difference is due to uncertainty in the ECMWF forecast. It is interesting to note that soundings for the 87% of the cases for which the Stratospheric Temperature Test was passed are of relatively high quality throughout the troposphere as well, with an RMS difference from ECMWF on the order of 1.7K in the lowest 1 km of the

atmosphere. This shows that the cloud clearing methodology works well in up to 90% cloud cover. Nevertheless, the accuracy of all these soundings is not considered high enough for either data assimilation or climate purposes. There is significant further improvement in tropospheric temperature profile accuracy, compared to that for cases passing the tropospheric temperature profile tests, using the smaller subset of cases passing the standard SST Test (21.45% of the ocean cases) but relatively little further improvement in those cases passing the Tight SST Test (9.12% of the cases), or the additional clear test (3.69% of the cases). For data assimilation purposes, we recommend experiments assimilating temperature profiles passing only the standard SST Test, on the one hand, and passing the test for the appropriate for the level of the temperature on the other hand, to assess the trade-off between coverage and accuracy. One might also consider assimilating lower tropospheric temperatures in cases passing the Mid-Tropospheric Temperature Test over ocean to further increase the spatial coverage of the data being assimilated.

Figure 5b shows analogous results for global accepted retrievals, including cases passing the Stratospheric Temperature Test, the Mid and Lower-Tropospheric Temperature Tests, and the clear test (which, over land, ice, and coasts, must also pass the Lower Troposphere Temperature Test). Error statistics in the stratosphere are again similar for cases passing the Stratospheric Temperature Test (81.91% of all cases) and either of the tropospheric tests (52.29% and 28.77%). Global agreement with ECMWF is slightly poorer than over ocean. A much larger difference in agreement with ECMWF occurs between all cases passing the Lower Tropospheric Temperature Test and the Mid-Tropospheric Temperature Test than over ocean, especially in the lower troposphere. For data assimilation purposes, we feel lower tropospheric temperatures retrieved over land should not be used when the Lower Tropospheric Temperature Test is not passed. Globally, 1.85% of the cases

passed the clear test, most of which were over non-frozen ocean. Retrievals in these cases are very accurate, but the global spatial coverage is very poor.

Figures 6a and 6b are analogous to Figures 5a and 5b but show statistics only for cases at a given pressure level passing the appropriate quality test. Statistics for cases passing the clear test (identical to those shown in Figures 5a and 5b) are included for comparison. Also shown are the accuracy of the regression first guess temperature profiles for all accepted retrievals and under clear conditions. The accuracy of the physical retrieval is higher than the regression, and more so under cloudy conditions than clear conditions. Part of this is due to the increased accuracy of $\hat{R}_i^4$, used to derive the final temperature profile, compared to $\hat{R}_i^1$, used to derive the regression first guess.

Figure 7 shows the RMS difference between retrieved 1 km tropospheric layer mean temperatures and the collocated ECMWF 3 hour forecast for all accepted cases as a function of retrieved effective cloud fraction. Results are shown for each of the 8 lowest 1 km layers of the atmosphere. Only those cases passing the appropriate temperature profile test are included in the statistics. Agreement degrades with increasing cloud cover, but only very slowly. The largest errors are in the 2 lowest layers in the atmosphere, at moderate to high cloud fraction, where the percentage acceptance rate is low. RMS temperature differences from ECMWF below 600 mb are somewhat larger than the 1 K goal for retrieval accuracy. Part of this difference can be attributed to the fact that the ECMWF forecast is not perfect. It is also possible that the accuracy of the ECMWF forecast may be somewhat poorer with increasing cloud cover.

Figures 8a and 8b are analogous to Figures 5a and 5b, but show RMS percent difference of retrieved 1 km layer precipitable water from the ECMWF "truth". In these, and other water vapor statistics, the percent error in a given case is weighted by the "truth" so as not to inflate percent

errors for very dry cases. These statistics should be used with caution, especially in the mid-upper troposphere, where considerable errors could exist in the ECMWF "truth". Nevertheless, over ocean, statistics are not appreciably different for cases passing the different tropospheric and ocean skin temperature thresholds. As with regard to temperature, a larger degradation occurs in agreement of humidity profile with ECMWF in the mid-lower troposphere over land when the looser constituent profile criteria are used. We recommend at this time to use the appropriate temperature test when attempting to assimilate water vapor at a given level of the atmosphere. Soundings passing either tropospheric temperature test also pass the constituent profile test because the temperature profile criteria are equal to, or tighter than, those in the constituent profile test. For climate purposes, it is better to include all cases passing the Constituent Test, so as to minimize a dry bias in the sample. Susskind et al. (2005) evaluate monthly mean interannual differences of water vapor and ozone fields using all cases passing the Constituent Test.

Figures 9a,b are analogous to Figures 6a,b and show water vapor % differences from "truth" for clear cases and cases passing the temperature test for the appropriate level. There is little apparent difference in water vapor retrieval occurring between clear cases and all cases passing the appropriate temperature profile test. While the regression first guess appears to be more accurate in the mid troposphere than the final retrieval, this may be an artifact because the regression was "trained" against ECMWF.

Figure 10 is analogous to Figure 7, but for percent differences from ECMWF of 1 km layer precipitable water as a function of retrieved effective fractional cloud cover. Only soundings passing the appropriate temperature profile test for a given level of the atmosphere (Mid-Tropospheric Temperature Test or Lower Tropospheric Temperature Test) are included in the statistics, as was done in Figure 9. Agreement with ECMWF degrades slightly with increasing

cloud cover in the lowest 2 km of the atmosphere, but not appreciably. Part of this could be due to sampling differences, because the AIRS retrievals determine water vapor in the clear portions of the partially cloudy scene, while the forecast values are for the whole scene. Differences from ECMWF above 500 mb are large and show no cloud fraction dependence. This may be a spurious result in which errors in ECMWF upper tropospheric water vapor are dominating the statistics for all cloud fractions.

The fundamental parameter used in the determination of geophysical parameters from AIRS/AMSU data is the clear column radiance $\hat{R}_i$, which represents the radiance AIRS channel i "would have seen" if no clouds were in the field of view. Geophysical parameters are determined which are consistent with $\hat{R}_i$. Derived geophysical parameters whose accuracy degrades slowly with increasing cloud cover implies that the accuracy of $\hat{R}_i$ also degrades slowly with increasing cloud cover. $\hat{R}_i$ is an important geophysical parameter derived from AIRS in its own right.

Figure 11a shows the mean value of $\hat{R}_i$ (in brightness temperature units) from 650 cm^{-1} to 756 cm^{-1} for all non-frozen ocean cases 50°N – 50°S on September 6, 2002 passing the Tight SST Test. The most opaque portion of the spectrum is near 667.5 cm^{-1}, and is primarily sensitive to atmospheric temperatures near 1 mb (50 km). Radiances in the surrounding spectral region are also primarily sensitive only to stratospheric temperatures and are not affected by clouds in the field of view. Radiances at frequencies greater than 690 cm^{-1} see increasing amounts of the troposphere, especially between absorption lines (the locally higher brightness temperatures) and are increasingly affected by cloud cover with increasing frequency. Radiances between lines at frequencies higher than 740 cm^{-1} are also increasingly sensitive to contributions from the ocean surface.

Figures 11b and 11c show the mean and standard deviation of the (tuned) differences between $\hat{R}_i$ and R_i computed from the "truth" for all cases in this geographic domain passing the Tight SST Test, the standard SST Test, the Lower Tropospheric Temperature Test, and the Mid-Tropospheric Temperature Test, respectively. Figure 11c also contains the channel noise spectrum. In this calculation, the "truth" is taken as the ECMWF forecast of temperature-moisture-ozone profile, along with the ECMWF ocean surface skin temperature. The Masuda Ocean surface emissivity model (1988), revised by Wu and Smith (1997), was used to generate the ocean surface emissivities in the calculation of the expected true radiances, assuming a surface wind speed of 5 m/sec. The surface contribution is the biggest uncertainty in the computation of the "truth" radiances because of errors in both the true ocean skin temperature and in the true surface emissivity.

It is apparent that the difference of clear column radiances from those computed from the truth increases only slightly in the more difficult cloud cases, and in general matches expected radiances to within the AIRS noise level. Standard deviations of observed minus computed brightness temperatures for stratospheric sounding channels are actually lower than the channel noise, because radiances of a AIRS fields of view are averaged together to produce the cloud cleared radiances. The increasing difference of clear column radiances from those computed from the "truth" between absorption lines above 740 cm^{-1} has a large component arising from errors in the "truth".

It is noteworthy that the biases of observed minus computed brightness temperatures are essentially zero for all cases, with some small negative biases between absorption lines at the frequencies sensitive to the lowest portions of the atmosphere in cases passing the Mid-Tropospheric Temperature Test as a result of small cloud clearing errors in these cases. First of all,

this implies that the tuning coefficients derived from clear ocean night ocean cases on September 6, 2002 are equally well applicable to a much larger ensemble of ocean cases on the same day. Secondly, it demonstrates that clear column radiances for cases passing the Mid-Tropospheric Temperature Test are essentially unbiased at most sounding channel frequencies. The standard deviations of the clear column radiances from "truth" are also only slightly dependent on the degree of cloud contamination. Errors in the "truth" dominate the standard deviations shown in Figure 11c, especially at 667.5 cm^{-1}, which is primarily sensitive to 1 mb temperature, and at frequencies sensitive to the ocean surface. In addition, the larger standard deviation at 679.31 cm^{-1} is a result of significant absorption by O_3, and those at 729.0 cm^{-1}, 730.8 cm^{-1}, and 745.1 cm^{-1}, and 754.4 cm^{-1} result from significant absorption by H_2O.

Figure 12 shows histograms of the difference between observed and computed brightness temperatures for the two channels indicated by the black dots in Figure 9, at 724.52 cm^{-1} and 749.19 cm^{-1} respectively. These frequencies are primarily sensitive to temperatures at 580 mb and 900 mb respectively, with a large surface contribution at 749.19 cm^{-1}. Results are shown for the four most stringent quality tests. The differences between the accuracy of clear column radiances at 724.52 cm^{-1}, for cases passing the different quality tests with spatial coverage ranging from 9.14% to 58.27%, are miniscule, with essentially no outliers in any category. Differences are somewhat larger at 749.19 cm^{-1}, but increase only slightly for cases passing the Mid-Tropospheric Temperature Test. For this reason, all clear column radiances are flagged as good for those cases passing the Mid-Tropospheric Temperature Test.

It is apparent from Figure 11 that the tuning coefficients derived for clear ocean night cases on September 6, 2002 are applicable to all ocean night cases on that day. Figures 13a-c show analogous results for all (global) cases passing the Mid-Tropospheric Temperature Test on

September 6, 2002 and January 25, 2003 corresponding to a different season and year. The biases (necessary tuning) are shown to be globally applicable, and also constant in time. Standard deviations from the truth at channels more sensitive to the surface are somewhat larger than for the non-frozen ocean cases because of larger errors in the "truth" arising from greater uncertainty in both surface skin temperature and spectral emissivity.

Operational numerical weather prediction centers currently assimilate radiance observations from IR sounders only for those cases thought to be unaffected by clouds (McNally, et al. 2000). This criterion severely limits the number of IR channel radiances being used in the assimilation processes, and tends to minimize the potential improvement in forecast skill achievable from optimal use of AIRS radiance observations. We encourage operational centers to attempt to use AIRS derived clear column radiances in their assimilation, applying the same quality control so as to accept only those clear column radiances "thought to be unaffected by clouds."

5. Summary

The AIRS Science Team algorithm used to analyze AIRS/AMSU data to derive geophysical parameters, including cloud cleared radiances, have been described. This methodology is essentially the same as that developed by Susskind et al. (2003) based on experience with simulated data. This algorithm is called Version 4. The major modifications involve treating the effects of errors in the Radiative Transfer Algorithm and a new concept of geophysical parameter dependent quality flags. Accurate soundings of temperature and moisture profiles, and corresponding clear column radiances, are produced in up to 90% effective fractional cloud cover. Accuracy of results, as judged by agreement with the ECMWF forecast, degrade only slightly with increasing cloud fraction when appropriate quality control is applied. The percent of accepted geophysical

parameters does decrease with increasing cloud fraction, in a manner which depends on the geophysical parameter.

The Goddard DAAC began analyzing near real time AIRS/AMSU data, using the Version 4 algorithm described in this paper, on April 1, 2005. The DAAC also began analysis of historical AIRS/AMSU data going back to September 1, 2002, using the Version 4 algorithm. Level 1B (radiances), Level 2 (spot by spot retrievals) and Level 3 (gridded) data is available. The Level 3 data is given on a 1° x 1° latitude-longitude grid, and averaged in 1 day, 8 day, and monthly mean segments, with ascending (1:30 PM local time) and descending (1:30 AM local time) orbits gridded separately. Examples of 1 day gridded ascending orbits are shown in this paper. Susskind et al. (2005) show examples of monthly mean gridded fields retrieved from AIRS/AMSU observations using the Version 4 algorithm. The data can be ordered at http://daac.gsfc.nasa.gov/data/datapool/AIRS/index.html. Collection 003 should be requested which has results derived using the Version 4 algorithm described in this paper. Collection 002 has results derived using an earlier algorithm (Version 3). These should not be used for scientific studies.

Research is continuing to improve the AIRS/AMSU retrieval algorithm. Version 5 should be made operational and delivered to the Goddard DAAC in 2006. Correcting the current limitations in retrieved land surface spectral emissivity and improving the quality of the error estimates of retrieved geophysical parameters are the highest priorities for the Version 5 algorithm. We also plan to use the improved error estimates directly as quality control indications, rather than using thresholds based on a number of different tests as done in Version 4. Research is continuing to optimize all other aspects of the retrieved algorithm as well.

References

Aumann, H., M. T. Chahine, C. Gautier, M. Goldberg, E. Kalnay, L. McMillin, H. Revercomb, P. Rosenkranz, W. Smith, D. Staelin, L. Strow, and J. Susskind, AIRS/AMSU/HSB on the Aqua mission: Design, science objectives, data products, and processing systems. *IEEE Trans. Geosci. Remote Sensing*, **41**, 253-264, February 2003.

Goldberg, M. D., Y. Qu, L. M. McMillin, W. Wolff, L. Zhou, and M. Divakarla, AIRS near-real-time products and algorithms in support of operational numerical weather prediction. *IEEE Trans. Geosci. Remote Sensing*, **41**, 379-389, February 2003.

Goldberg, M., C. Barnet, and M. Divakarla, Three-way comparisons of AIRS, ATOVS and radiosonde temperature and moisture profiles. Submitted to *JGR*, 2005.

Masuda, K., T. Takashima, and Y. Takayama, Emissivity of pure and sea waters for the model sea surface in the infrared window region. *Remote Sensing of the Environ.*, **24**, 313-329, 1988.

McNally, A., J. Derber, W. Wu, and B. Katz, The use of TOVS level 1B radiances in the NCEP SSI Analysis system. *Q.J.R. Meteorol. Soc.*, **126**, 689-724, 2000.

Rosenkranz, P. W., Retrieval of temperature and moisture profiles from AMSU-A and AMSU-B measurements. In *Proc. IGARSS*, 2000.

Strow, L., S. Hannon, S. De-Souza Machado, and D. Tobin, Validation of the AIRS radiative transfer algorithm. Submitted to *JGR*, 2005.

Susskind and Pfaendtner, Impact of interactive physical retrievals on NWP. Proceedings of a workshop held at ECMWF, May 9-12, 1989, European Center for Medium-Range Weather Forecasts, Shinfield park, Reading TG2 9AX, United Kingdom, 1989.

Susskind, J., P. Piraino, L. Rokke, L. Iredell, and A. Mehta, Characteristics of the TOVS Pathfinder Path A dataset. *Bull. Amer. Meteor. Soc.*, **78**, No. 7, 1449-1472, 1997.

Susskind, J., C. D. Barnet, and J. M. Blaisdell, Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds. *IEEE Trans. Geosci. Remote Sensing*, **41**, 390-409, February 2003.

Susskind, J. and R. Atlas, Atmospheric soundings from AIRS/AMSU/HSB. *Proc. SPIE Conference*, **5425-31**, 311-319, April 12-15, 2004.

Susskind, J., L. Iredell, F. Keita, and G. Molnar, Validatin of interannual differences of AIRS monthly mean parameters. Submitted to *JGR*, 2005.

Wu, X. and W. L. Smith, Emissivity of rough sea surface for 8-13 μm: Modeling and verification, *Appl. Opt.*, **36**, 2609-2619, 1997.

Figure Captions

Figure 1. Percent accepted retrievals as a function of retrieved effective cloud fraction for different acceptance tests. The average cloud fraction for all cases in indicated, as well as for all accepted cases for each test.

Figure 2. Values and spatial coverage of cloud parameters, 200 mb temperature, total precipitable water, and 500 mb temperature for ascending orbits on September 6, 2002. Gray indicates missing data, or areas where retrievals fail the appropriate quality test.

Figure 3. Differences from ECMWF "truth", and spatial coverage using different test thresholds, for both 700 mb temperature and sea surface temperature.

Figure 4. Histogram of number of accepted cases as a function of differences of retrieved sea surface temperatures from ECMWF "truth" for different quality tests. Also indicated are the mean differences, the standard deviation, the percent of the cases passing the test, and the percent of cases differing by more than 3K from the truth.

Figure 5a. RMS differences of retrieved layer mean temperatures from ECMWF "truth" for all ocean 50°N-50°S cases passing different quality tests. The percent of cases accepted, relative to all possible cases, is indicated for each test, as well as the total accepted cases.

Figure 5b. As in Figure 5a, but for global cases.

Figure 6a,6b. As in Figures 5a, 5b, but each level includes all cases passing the appropriate temperature profile test. RMS errors of the regression first guess are also included in the figure. The solid gray curves are identical to those in Figures 5a, 5b.

Figure 7. RMS 1 km layer mean temperature differences from the "truth" as a function of retrieved cloud fraction for all cases passing the appropriate temperature quality test.

Figure 8a,b. As in Figures 5a, 5b, but for percent differences of 1 km layer precipitable water from the ECMWF "truth".

Figure 9. As in Figure 6, but for percent differences in 1 km layer precipitable water.

Figure 10. As in Figure 7, but for percent differences of 1 km layer precipitable water.

Figure 11a. Mean retrieved clear column brightness temperatures for all non-frozen ocean 50°N-50°S cases on September 6, 2002 passing the Mid-Troposphere Temperature Test.

Figure 11b,c. Mean and standard deviation of the difference of retrieved clear column brightness temperatures from those computed from ECMWF "truth" for all cases passing different quality tests. AIRS channel noise is also indicated in Figure 11c.

Figure 12a,b. As in Figure 4 but for difference of retrieved clear column brightness temperatures for two channels indicated in Figure 11a.

Figure 13a-c. As in Figures 11a-c, but for global cases passing the Mid-Tropospheric Temperature Test for September 6, 2002 and January 25, 2003.

Table 1

QUALITY FLAG TEST THRESHOLDS

	Susskind et al. (2003)	Version 4					
Test	Acceptable Profile	1) T(p) good 200mb&up	2) q(p) good O_3(p) good	3) T(p) good 3km&up	4) T(p) good above surface	5) SST good Loose	6) SST good Tight
$\alpha\varepsilon$	80%	**90%**	90%	90%	90%	90%	90%
W_{liq}	.03	X	**.03**	.03	.03	**.01**	.01
$\Delta T(p)$	1.25	X	X	**2.0**	2.0	2.0	2.0
$A^{(4)}$	3	X	**8.0**	**2.0**	2.0	2.0	2.0
$A_{eff}^{(4)}$	8	X	X	X	**15** (X)	**8**	8
ΔF	1.75	X	**8.0**	**2.0 (6.0)**	**1.5 (1.5)**	1.5	1.5
R_{temp}	1.0	X	X	**0.75**	0.75	0.75	0.75
R_{surf}	1.0	X	X	**0.75** (X)	0.75 (X)	0.75	0.75
$A_{eff}^{(1)}$	X	**200**	200	**30** (X)	30 **(30)**	**9**	**5**
$\Delta\Theta_5$	X	X	X	**2.0**	2.0	2.0	2.0
Δ_{tskin}	X	X	X	X	**1.5**	1.5	1.5
RS	X	**10**	10	**4**	4	**1.2**	1.2

$\alpha\varepsilon$	is the effective cloud fraction
W_{liq}	is cloud liquid water
$\Delta T(p)$	represents the difference of retrieved lower tropospheric temperatures between MW only and IR/MW retrievals
$A^{(4)}$	represents the final channel noise amplification factor
$A_{eff}^{(4)}$	represents the final effective channel noise amplification factor
ΔF	represents the quality of the initial cloud clearing fit
R_{temp}	represents the degree to which the final temperature profile retrieval has converged
R_{surf}	represents the degree to which the final surface parameter retrieval has converged
$A_{eff}^{(1)}$	represents the initial effective channel noise amplification factor
$\Delta\Theta_5$	represents the agreement between the observed AMSU channel 5 brightness temperature and that computed from the final solution
Δ_{tskin}	represents the difference between the final surface skin temperature and the regression valve
RS	represents the principle component reconstruction score of the observed AIRS radiances

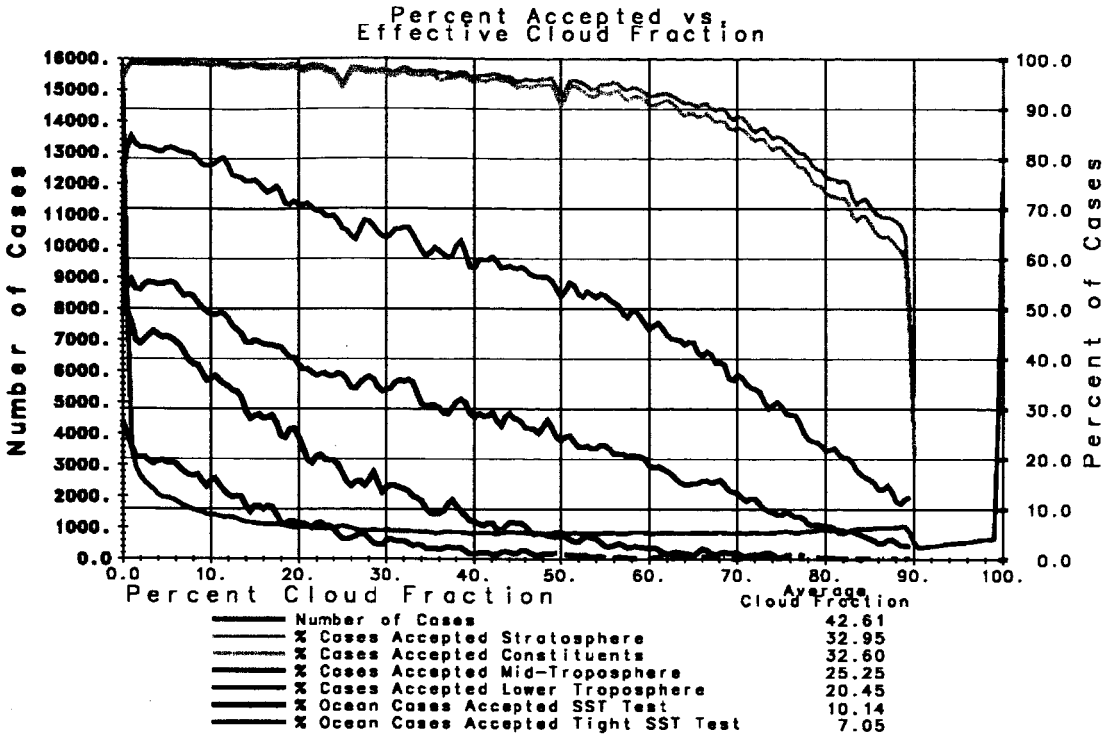

Figure 1

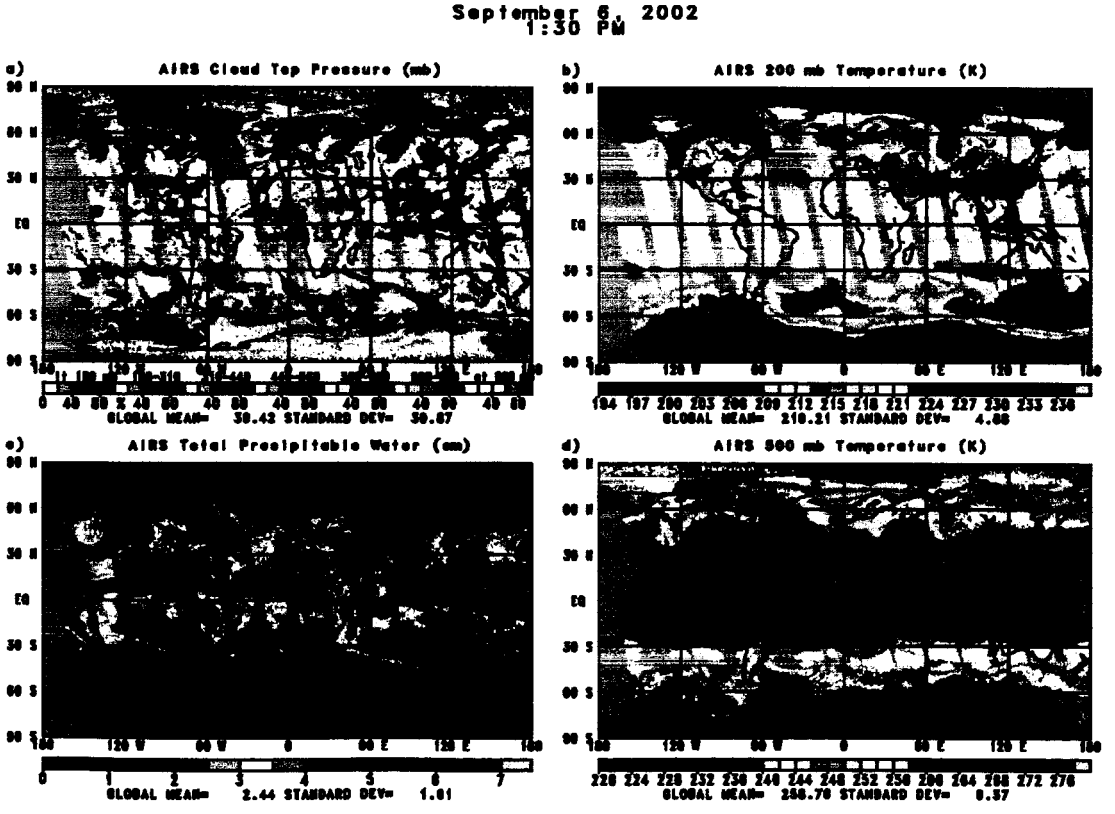

Figure 2

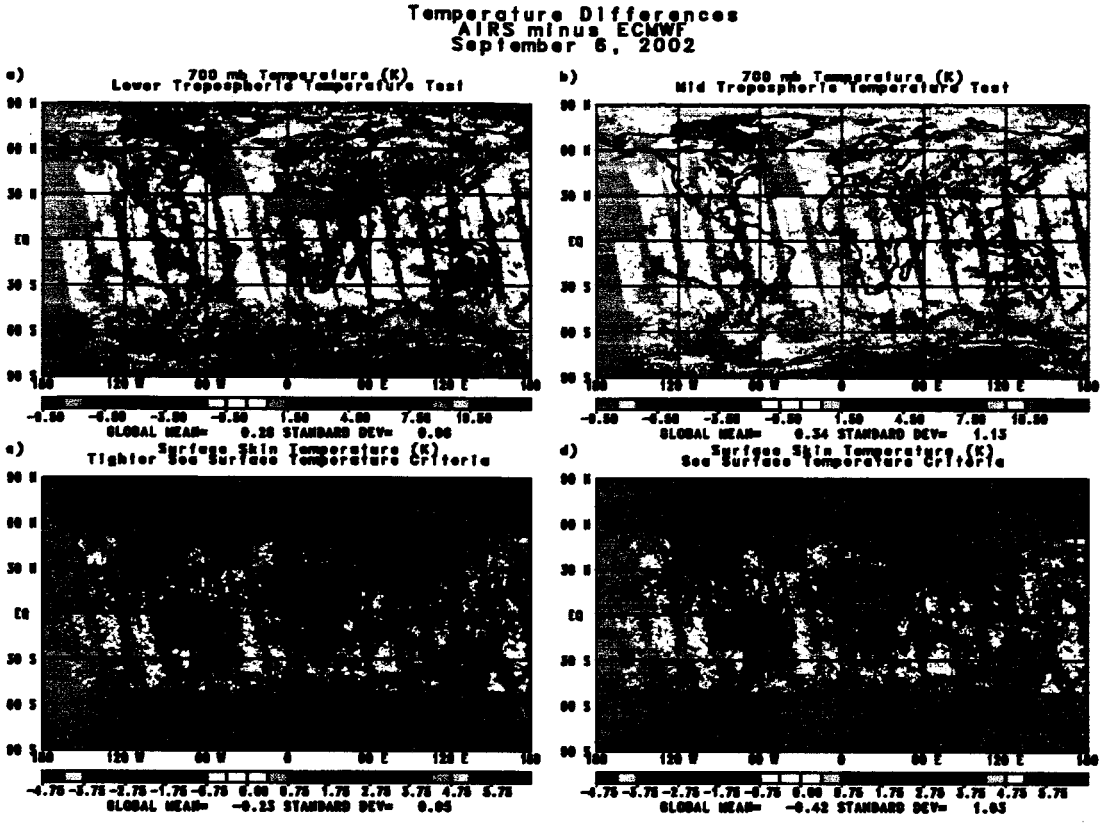

Figure 3

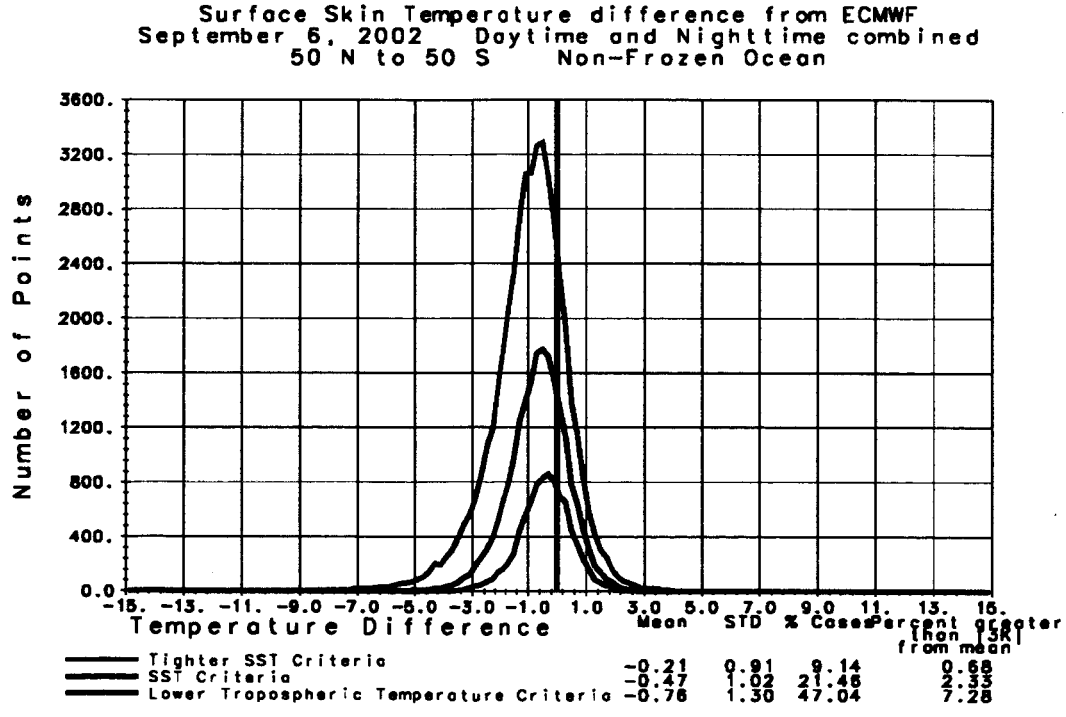

Figure 4

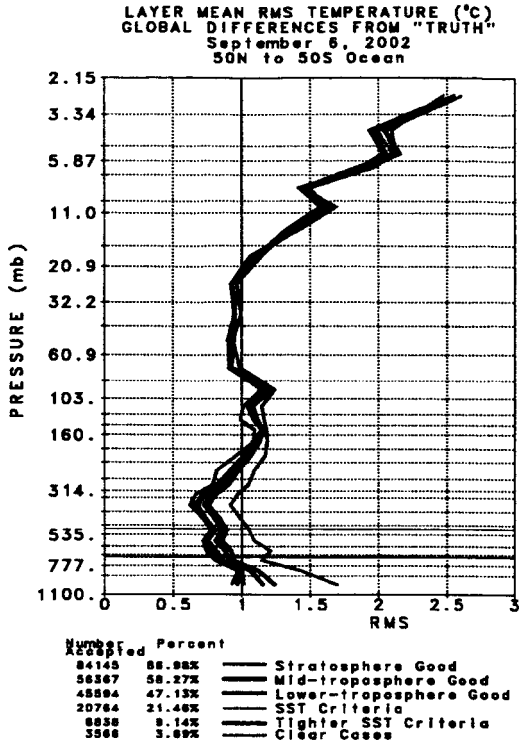

Figure 5a

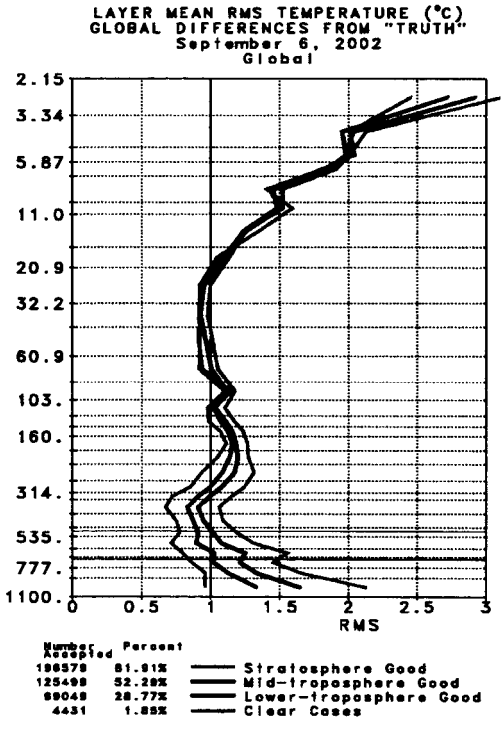

Figure 5b

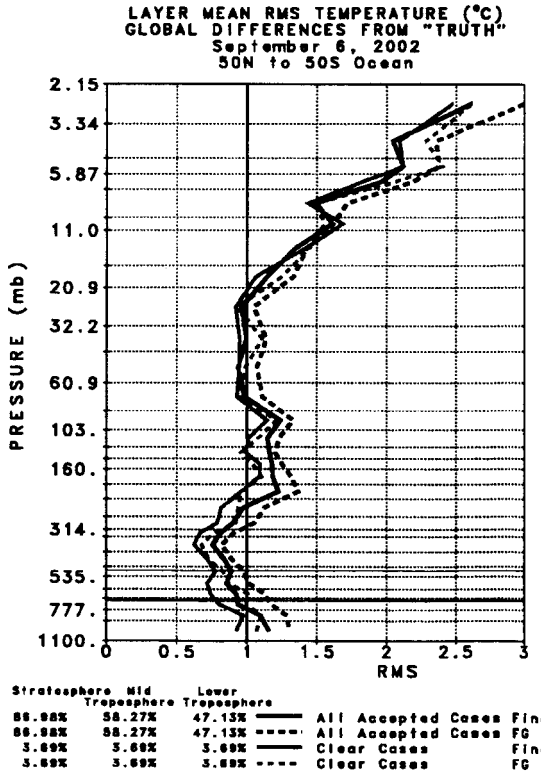

Figure 6a

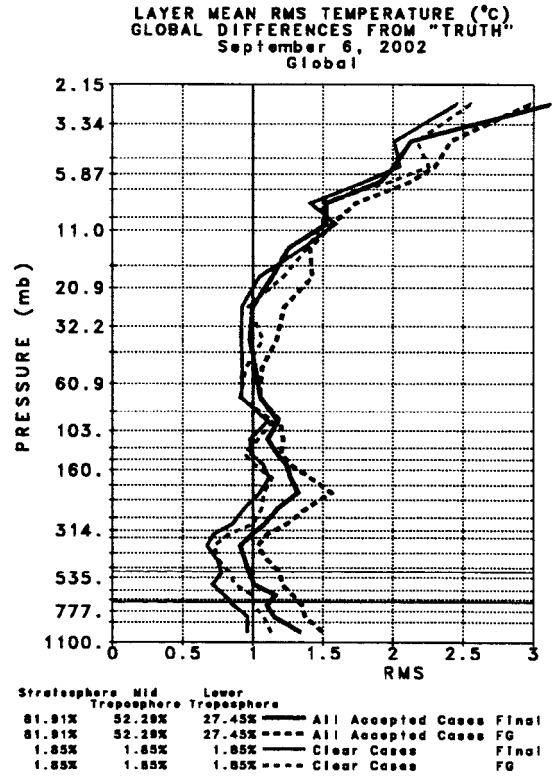

Figure 6b

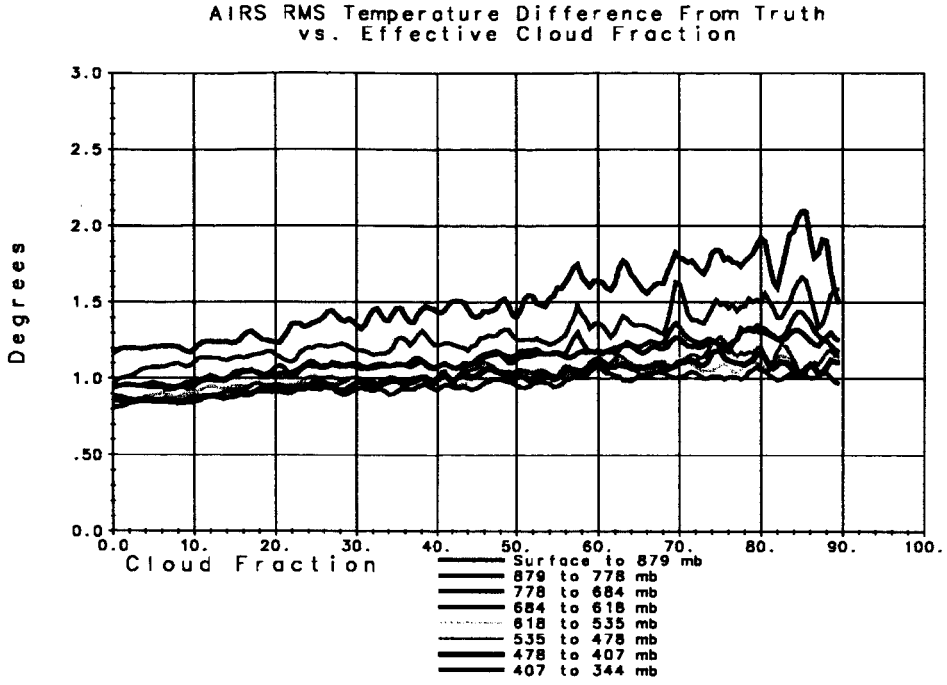

Figure 7

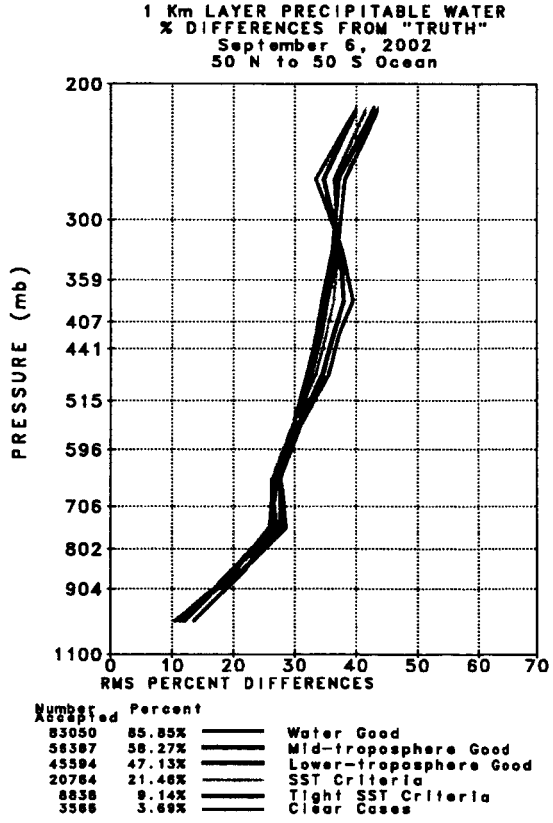

Figure 8a

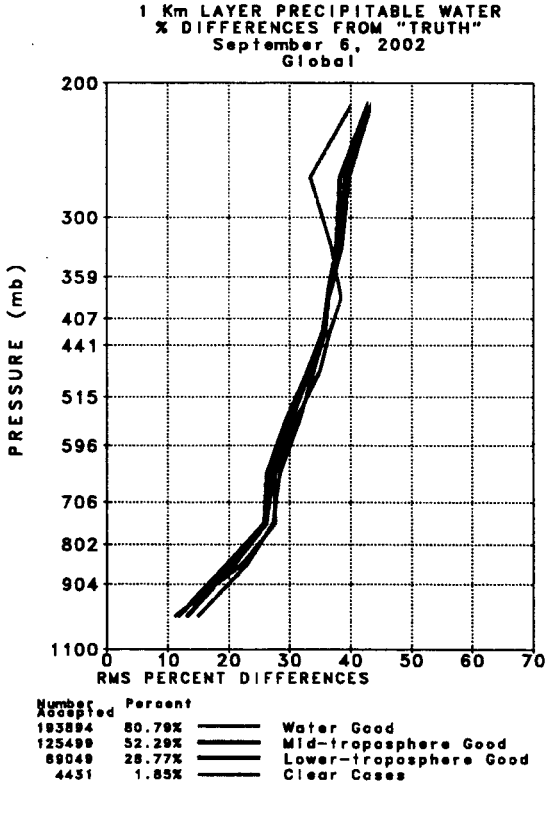

Figure 8b

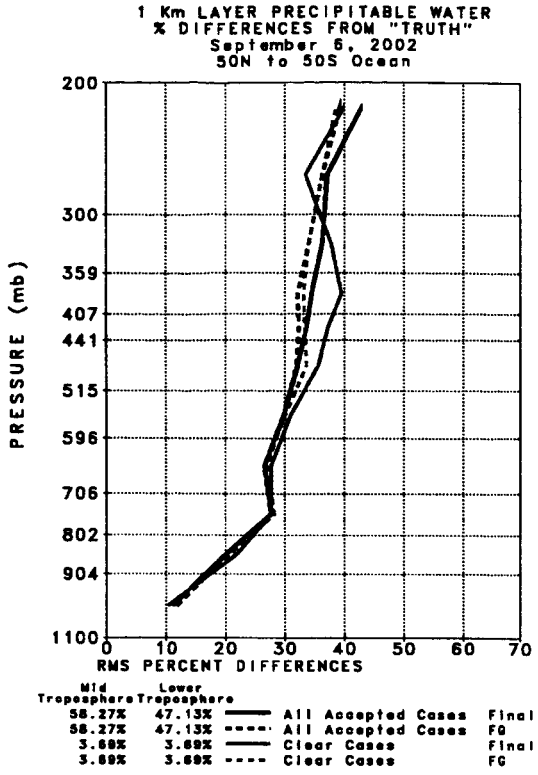

Figure 9a

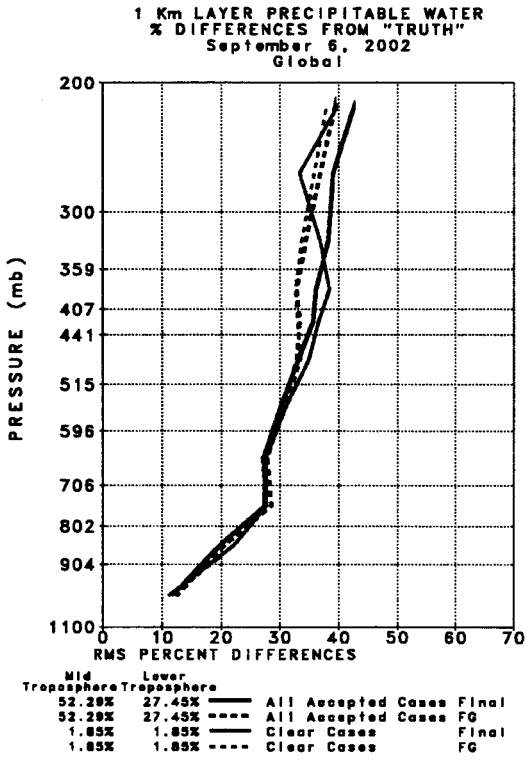

Figure 9b

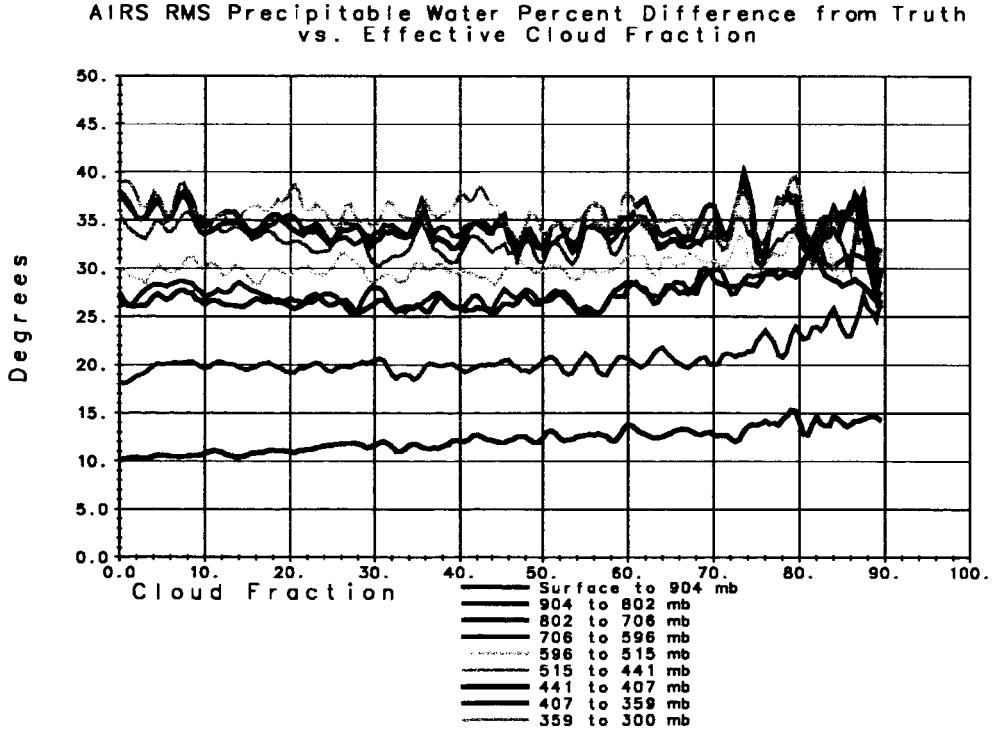

Figure 10

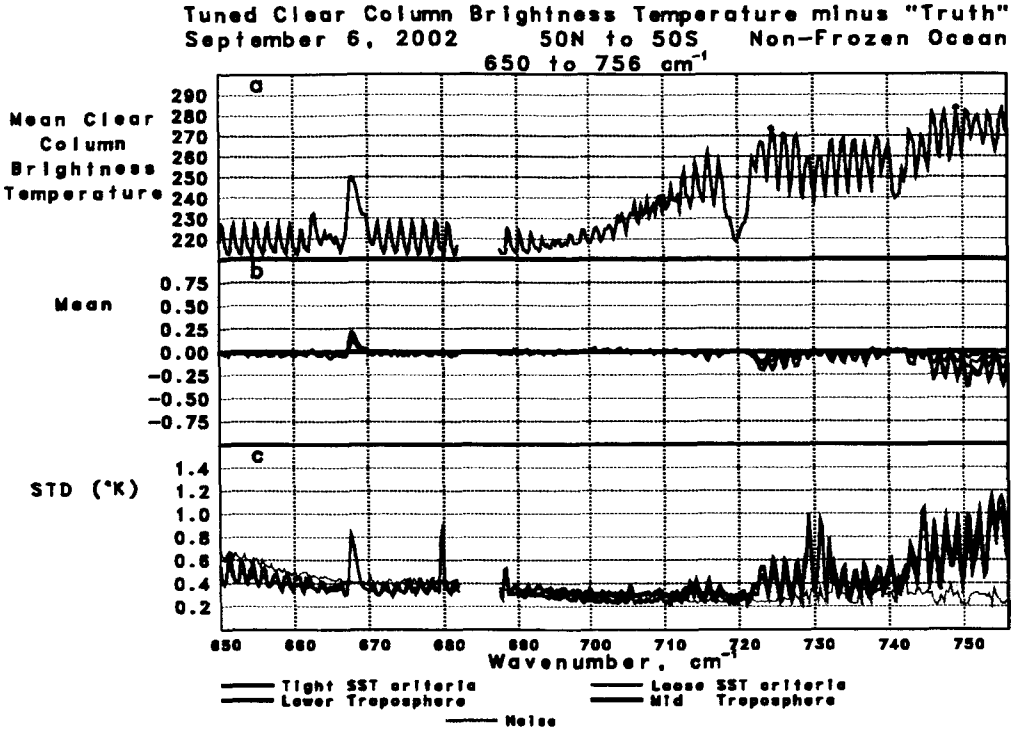

Figure 11

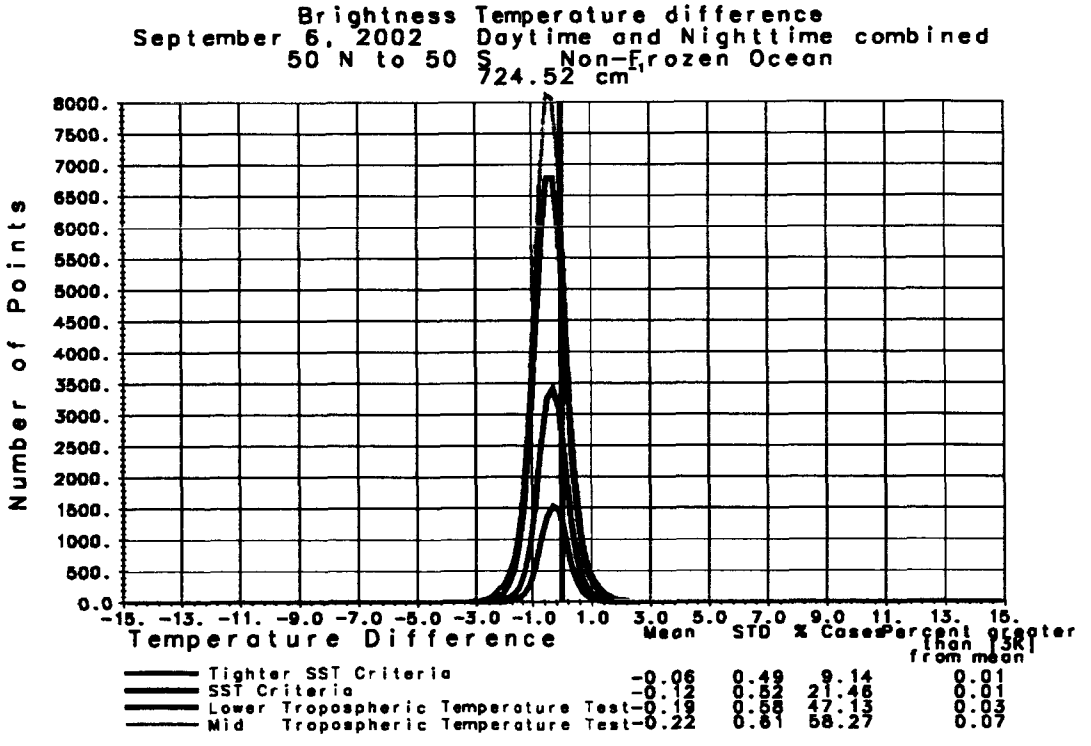

Figure 12a

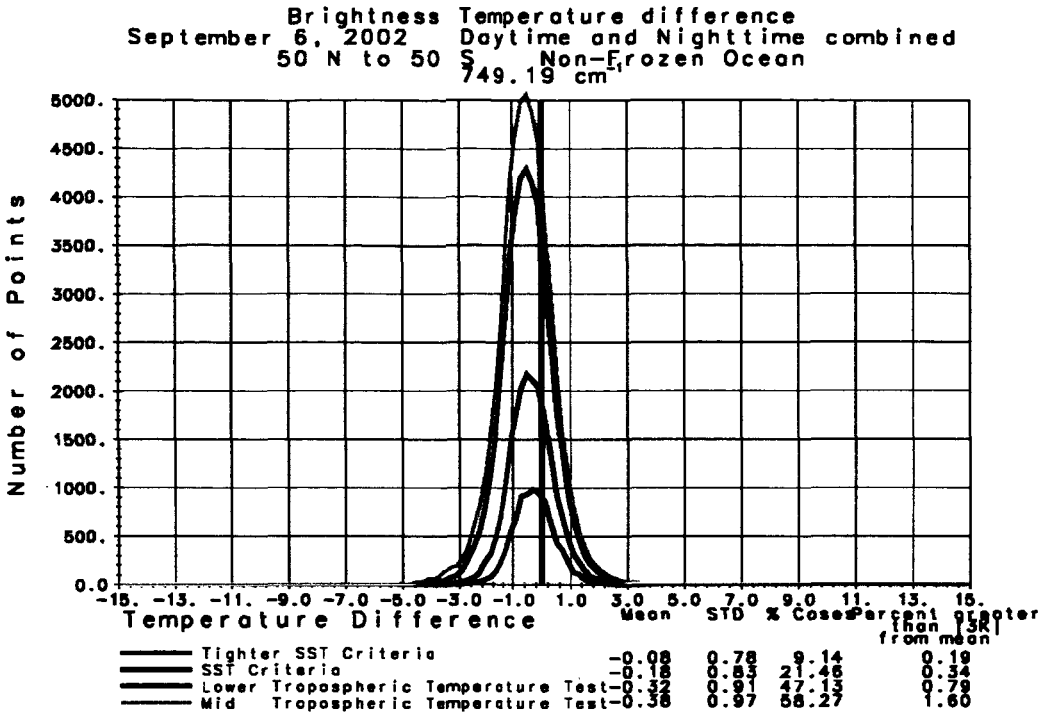

Figure 12b

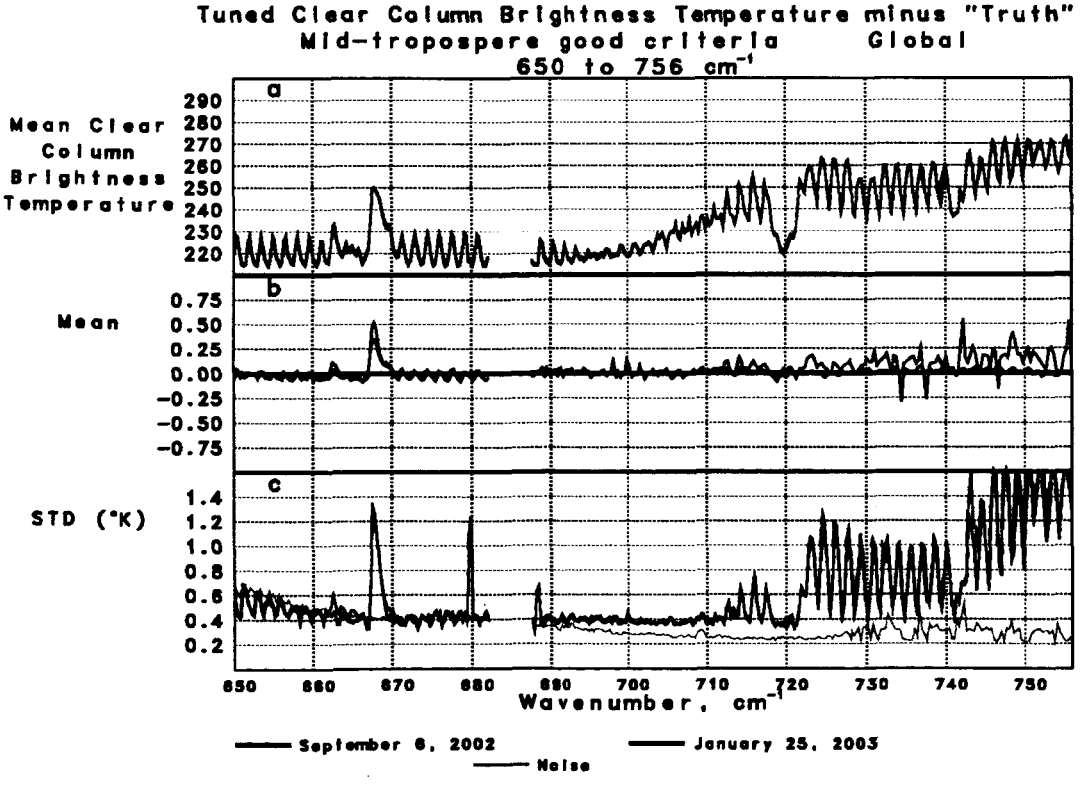

Figure 13